

SCORE Search Results Details for Application 10573229 and Search Result 20090528_121059_us-10-573-229a-1.rst.

Score Home	Retrieve Application	SCORE System	SCORE	Comments /
Page	List	Overview	FAQ	Suggestions

This page gives you Search Results detail for the Application 10573229 and Search Result 20090528_121059_us-10-573-229a-1.rst.

[Go Back to previous page](#)

GenCore version 6.3
Copyright (c) 1993 - 2009 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 31, 2009, 21:51:55 ; Search time 1876 Seconds
(without alignments)
47444.276 Million cell updates/sec

Title: US-10-573-229A-1
Perfect score: 920
Sequence: 1 tctgtagaggggaatggctg.....acccccaaagaaaccttcta 920

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 83780570 seqs, 48372533981 residues

Total number of hits satisfying chosen parameters: 167561140

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*

9: gb_est9:*
 10: gb_est10:*
 11: gb_est11:*
 12: gb_est12:*
 13: gb_est13:*
 14: gb_est14:*
 15: gb_est15:*
 16: gb_est16:*
 17: gb_est17:*
 18: gb_est18:*
 19: gb_gss1:*
 20: gb_gss2:*
 21: gb_gss3:*
 22: gb_gss4:*
 23: gb_gss5:*
 24: gb_gss6:*
 25: gb_gss7:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result			%	Query			ID	Description
	No.	Score	Match	Length	DB			
	1	920	100.0	920	5	BU183861		BU183861 AGENCOURT
	2	149.8	16.3	988	3	BE733157		BE733157 601567451
	3	108	11.7	533	10	DB059978		DB059978 DB059978
c	4	104.8	11.4	478	2	AW269751		AW269751 xv36g11.x
c	5	104.8	11.4	484	2	BE046932		BE046932 hd92f06.x
	6	104.4	11.3	500	10	DB100179		DB100179 DB100179
c	7	103.2	11.2	757	20	CC492937		CC492937 CH240_327
	8	101.4	11.0	526	10	DB073121		DB073121 DB073121
	9	101.2	11.0	669	22	EI735387		EI735387 CHORI5142
c	10	101.2	11.0	702	4	BQ021722		BQ021722 UI-H-DH1-
	11	95.4	10.4	549	10	DB074846		DB074846 DB074846
c	12	94.8	10.3	446	2	AW297623		AW297623 UI-H-BW0-
c	13	94.6	10.3	533	1	AA535137		AA535137 nf87h12.s
	14	94.6	10.3	637	2	AW970418		AW970418 EST382499
c	15	94.2	10.2	429	20	CE539857		CE539857 tigr-gss-
	16	94	10.2	707	22	EI399423		EI399423 MUGQ_CH25
	17	94	10.2	805	22	ED482970		ED482970 MUGQ_CH25
c	18	92.8	10.1	381	25	CU051182		CU051182 Equus cab
c	19	92.8	10.1	466	25	CU288702		CU288702 Equus cab
c	20	92.4	10.0	548	10	DB349144		DB349144 DB349144
c	21	91.8	10.0	454	4	BM667926		BM667926 UI-E-DW0-
	22	91.8	10.0	503	4	BM696584		BM696584 UI-E-DW0-

	23	91.8	10.0	1147	3	BG119504	BG119504	602349294
	24	91.4	9.9	449	1	AA426503	AA426503	zw02b10.r
c	25	91.4	9.9	682	6	CA431350	CA431350	UI-H-FG1-
	26	91.4	9.9	688	5	BX104512	BX104512	BX104512
	27	91.4	9.9	830	3	BG678891	BG678891	602624794
c	28	91	9.9	703	25	CU177831	CU177831	Equus cab
c	29	90.8	9.9	284	3	BF591952	BF591952	7o22d02.x
c	30	90.8	9.9	365	3	BF476176	BF476176	naa29d02.
c	31	90.8	9.9	421	1	AI767873	AI767873	wi97h11.x
c	32	90.8	9.9	444	1	AI218734	AI218734	oo07b06.x
c	33	90.8	9.9	448	3	BF516187	BF516187	UI-H-BW1-
c	34	90.8	9.9	483	1	AI697816	AI697816	we17f06.x
c	35	90.8	9.9	493	1	AI670048	AI670048	we64d02.x
c	36	90.8	9.9	508	1	AI668692	AI668692	zb63a11.x
c	37	90.8	9.9	625	1	AI990414	AI990414	wt74a11.x
c	38	90.8	9.9	629	5	BU730892	BU730892	UI-E-CI1-
c	39	89	9.7	755	24	AG621567	AG621567	Macaca fu
c	40	88.6	9.6	834	20	CC062054	CC062054	MUGQ_CH25
	41	88.4	9.6	496	25	CU026287	CU026287	Equus cab
	42	87.6	9.5	457	10	DB489743	DB489743	DB489743
c	43	87.4	9.5	349	18	M85416	M85416	EST01931 Fe
	44	87.4	9.5	760	24	AG671147	AG671147	Macaca fu
c	45	87.2	9.5	739	25	CT974461	CT974461	Equus cab

ALIGNMENTS

RESULT 1

BU183861

LOCUS BU183861 920 bp mRNA linear EST 04-SEP-2002

DEFINITION AGENCOURT_7964945 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6162433
5', mRNA sequence.

ACCESSION BU183861

VERSION BU183861.1 GI:22697845

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 920)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC/DCTD/DTP

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM13516 row: f column: 02
 High quality sequence stop: 633.

FEATURES Location/Qualifiers
 source 1. .920
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6162433"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_72"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 2 kb. Library constructed by Life
 Technologies."

ORIGIN

Query Match 100.0%; Score 920; DB 5; Length 920;
 Best Local Similarity 100.0%; Pred. No. 1.6e-249;
 Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	TCTGTAGAGGGGAATGGCTGCTGTGTCATGGGGGTGCATGAGCAGCCCAGTGGAGAGGTG	60
Db	1	TCTGTAGAGGGGAATGGCTGCTGTGTCATGGGGGTGCATGAGCAGCCCAGTGGAGAGGTG	60
Qy	61	CACTTGGTGAGAAACCGATGCCTCTGCCAACCACCTGCACTAACCTGCTGGGTCTGAGAC	120
Db	61	CACTTGGTGAGAAACCGATGCCTCTGCCAACCACCTGCACTAACCTGCTGGGTCTGAGAC	120
Qy	121	TGAGCCACTTTGGAAGCTGATCTTGAGCACCAGTCAAGCCCTTAGCTGGCTGCAGCCAC	180
Db	121	TGAGCCACTTTGGAAGCTGATCTTGAGCACCAGTCAAGCCCTTAGCTGGCTGCAGCCAC	180
Qy	181	AGCCAACAACAAGACTGCAACCTCCTGGGGGATCCTGAGCCAGAATCCCCTGGCTAAATT	240
Db	181	AGCCAACAACAAGACTGCAACCTCCTGGGGGATCCTGAGCCAGAATCCCCTGGCTAAATT	240
Qy	241	GCTCCTTGATTCTTAACCCACAGAAATTGTGTAAGACCTCCATCAGGTGTCGACAAGGAA	300
Db	241	GCTCCTTGATTCTTAACCCACAGAAATTGTGTAAGACCTCCATCAGGTGTCGACAAGGAA	300
Qy	301	GATCCCAGTAGGGCAGGAGACAGGAGCACCTCTGCTGTGGCCAATGCAGGAATGCTGGCC	360
Db	301	GATCCCAGTAGGGCAGGAGACAGGAGCACCTCTGCTGTGGCCAATGCAGGAATGCTGGCC	360

Qy	361	ATCATTGCTTCTGCTGGGCGACTGAGAAGCATCACCCACTTCCCCAGAACCTTTTTTTACG	420
Db	361	ATCATTGCTTCTGCTGGGCGACTGAGAAGCATCACCCACTTCCCCAGAACCTTTTTTTACG	420
Qy	421	TGGAGTGAAAACTTTAAGGGGCTGTCCAGCTAAACCTCCAACCTCCAGATCCCATGCCAA	480
Db	421	TGGAGTGAAAACTTTAAGGGGCTGTCCAGCTAAACCTCCAACCTCCAGATCCCATGCCAA	480
Qy	481	TTTCTCTGCTTCTGCAAAAGGACTTCAAGTGAAAGACATCTGCAGCTGTGAACGGGGGTA	540
Db	481	TTTCTCTGCTTCTGCAAAAGGACTTCAAGTGAAAGACATCTGCAGCTGTGAACGGGGGTA	540
Qy	541	AAACCCTCCCTGCCCCAGGCCCAAGCAAGGATTTCCCTAGCGGGGAGGAAGGTAGAATC	600
Db	541	AAACCCTCCCTGCCCCAGGCCCAAGCAAGGATTTCCCTAGCGGGGAGGAAGGTAGAATC	600
Qy	601	GAGAGACCTCTAACCCTGGGAGAGGAGGGAGGGAAATCTCCGAGGACCAGGGTTATGCAA	660
Db	601	GAGAGACCTCTAACCCTGGGAGAGGAGGGAGGGAAATCTCCGAGGACCAGGGTTATGCAA	660
Qy	661	CAACACAAGGGAAGTACCTGCTGGGTCTGGGGGTGGGGAAGGAAAATCCCTACTGCCC	720
Db	661	CAACACAAGGGAAGTACCTGCTGGGTCTGGGGGTGGGGAAGGAAAATCCCTACTGCCC	720
Qy	721	CAAGAGCCAGCCCCGAACCCAAGGCACAGCTTATACTGGCCCCGGGGCCTGGGGGGGCAC	780
Db	721	CAAGAGCCAGCCCCGAACCCAAGGCACAGCTTATACTGGCCCCGGGGCCTGGGGGGGCAC	780
Qy	781	GAAAACCTTGAAAAAGGGGCGCCTTCCCAGCTTCCCCGGGGGTAAAGGCTTTACCCCCCA	840
Db	781	GAAAACCTTGAAAAAGGGGCGCCTTCCCAGCTTCCCCGGGGGTAAAGGCTTTACCCCCCA	840
Qy	841	GAGGGGGGGGGGAAAAATCCGAGTGGGATCTTTCCCAACCGCCGAAGACTAAAACCTTTAA	900
Db	841	GAGGGGGGGGGGAAAAATCCGAGTGGGATCTTTCCCAACCGCCGAAGACTAAAACCTTTAA	900
Qy	901	ACCCCCAAAGAAACCTTCTA	920
Db	901	ACCCCCAAAGAAACCTTCTA	920

RESULT 2
BE733157
LOCUS BE733157 988 bp mRNA linear EST 15-SEP-2000
DEFINITION 601567451F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3842292 5',
mRNA sequence.
ACCESSION BE733157
VERSION BE733157.1 GI:10147149

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 988)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM535 row: e column: 13
High quality sequence stop: 703.

FEATURES Location/Qualifiers

source 1..988
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3842292"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_21"
/note="Organ: placenta; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 16.3%; Score 149.8; DB 3; Length 988;
Best Local Similarity 90.4%; Pred. No. 1.3e-30;
Matches 160; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 127 ACTTTGGAAGCTGATCTTGGAGCACCAGTCAAGCCCTTAGCTGGCTGCAGCCACAGCCAA 186
|||||

Db 1 ACTTTGGAAGCTGATCTTGGAGCACCAGTCAAGCCCTTAGCTGGCTGCAGCCACAGCCAA 60

Qy 187 CAACAAGACTGCAACCTCCTGGGGGATCCTGAGCCAGAATCCCCTGGCTAAATTGCTCCT 246
|||||

Db 61 CAACAAGACTGCAACCTCCTGGGGGATCCTGAGCCAGAATCCCCTGGCTAAATTGCTCCT 120

Qy 247 TGATTCTTAACCCACAGAAATTGTGTAAGACCTCCATCAGGTGTCGACAAGGAAGAT 303
||||| || | ||| | | |||| |

Db 121 TGATTCTTAACCCACAGAAATTGTGCTTAACACCATGCAGAAGCTGCCAAGGCTTAT 177

RESULT 3

DB059978

LOCUS DB059978 533 bp mRNA linear EST 18-JAN-2008

DEFINITION DB059978 TESTI2 Homo sapiens cDNA clone TESTI2053144 5', mRNA
sequence.

ACCESSION DB059978

VERSION DB059978.1 GI:83532935

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 533)

AUTHORS Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.TITLE Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes

JOURNAL Genome Res. 16 (1), 55-65 (2006)

PUBMED 16344560

COMMENT Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.

FEATURES Location/Qualifiers

source 1. .533
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TESTI2053144"

/tissue_type="testis"
/clone_lib="TESTI2"
/note="Vector: pME18SFL3"

ORIGIN

Query Match 11.7%; Score 108; DB 10; Length 533;
Best Local Similarity 67.5%; Pred. No. 8.3e-19;
Matches 181; Conservative 0; Mismatches 85; Indels 2; Gaps 2;

```
Qy      9  GGGGAATGGCTGCTGTGTGTCATGGGGGTGCATGAGCAGCCCAGTGGAGAGGTGCACTTGGT 68
      ||| |   ||| |   ||| || || |||   || ||||  ||||| ||| |||
Db     22  GGGAAGCAGCTACCATGTTGTGAGGCTGCTCAAGTAGCCTTGTGGAGAGGTCCACCTGGA 81

Qy     69  GAGAAACCGATGCCTCTGCCAACCACCTGCACTAACCTGCTGGGTCTGAGACTGAGCCAC 128
      ||||| | | |||| |||| ||  || |||| ||| |||   | || | |||||
Db     82  GAGAAAACAAGGCCTTTGCC-ACAGCCGGCACCAACTTGCCAACCATCTGAGTCAGCCAC 140

Qy    129  TTTGGAAGCTGATCTTGGAGCACCAGTCAAGCCCTTAGCTGGCTGCAGCCACAGCCAACA 188
      ||||| |||| | ||| || || |||| | || || | | | |||| | |
Db    141  CTTGGAAGTGGATCCTCCAGCCTCATTCAGCCTTCAGATGACCACTGTCCCAGCTCATA 200

Qy    189  ACAAGACTGCAACCTCCTGGGGGATCCTGAGCCAGAATCCCCTGGCTAAATTGCTCCTTG 248
      | ||||| ||||| || | || | ||||| || || |||| | || |||
Db    201  TCTAGACTGCAACCTCATGAGAGACCTGGAGCCAGAA-CACCCAGCTAAGCTGTTCTGA 259

Qy    249  ATTCTTAACCCACAGAAATTGTGTAAGA 276
      ||||  || ||||| | | || ||||
Db    260  ATTCCCGACTCACAGAGACTATGAAAGA 287
```

RESULT 4

AW269751/c

LOCUS AW269751 478 bp mRNA linear EST 03-JAN-2000

DEFINITION xv36g11.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone

IMAGE:2815268 3', mRNA sequence.

ACCESSION AW269751

VERSION AW269751.1 GI:6656781

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 478)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -40UP from Gibco

High quality sequence stop: 446.

```

FEATURES
    source          Location/Qualifiers
                    1. .478
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:2815268"
                    /lab_host="DH10B"
                    /clone_lib="Soares_NFL_T_GBC_S1"
                    /note="Organ: pooled; Vector: pT7T3D-PacI; Site_1: Not I;
                    Site_2: Eco RI; Equal amounts of plasmid DNA from three
                    normalized libraries (fetal lung NbHL19W, testis NHT, and
                    B-cell NCI_CGAP_GCB1) were mixed, and ss circles were made
                    in vitro. Following HAP purification, this DNA was used as
                    tracer in a subtractive hybridization reaction. The driver
                    was PCR-amplified cDNAs from pools of 5,000 clones made
                    from the same 3 libraries. The pools consisted of
                    I.M.A.G.E. clones 297480-302087, 682632-687239,
                    726408-728711, and 729096-731399. Subtraction by Bento
                    Soares and M. Fatima Bonaldo.  "

```

ORIGIN

```

Query Match          11.4%;  Score 104.8;  DB 2;  Length 478;
Best Local Similarity 68.5%;  Pred. No. 6.5e-18;
Matches 174;  Conservative 0;  Mismatches 77;  Indels 3;  Gaps 2;

```

```

Qy      24  TGTCATGGGGGTGCATGAGCAGCCCAGTGGAGAGGTGCACTTGGTGAGAAACCGATGCCT  83
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      340  TTTCATGAGGATACTCAAGCATTCTATGGAGAGATCCACATGGTGAGAAACTGAAGCCT  281

Qy      84  -CTGCCAACCACCTGCACTAACCTGCTGGGTCTGAGACTGAGCCACTTTGGAAGCTGATC  142
      | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      280  CCTACCAAGAGCCAGCACCAACTTGCCAGCTATGTGAATGAGCCATCTTAGAAGTGGGTT  221

Qy     143  TTGGAGCACCAAGTCAAGCCCTTAGCTGGCTGCAGCCACAGCCAACAACAAGACTGCAACC  202
      | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     220  CTCTAGCCCTAGTCAGGCCTTCATATGACTGCAGCCAGGGCTGATATTTTGACTACAACC  161

Qy     203  TCCTGGGGGATCCTGAGCCAGAATCCCCTGGCTAAATTGCTCCTTGATTCTTAACCCACA  262
      | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     160  TCATGAGAGA--CTGAGCCACAACAACCTAGCTAAGAAGCTCCTGAATTCCTACCAACA  103

Qy     263  GAAATTGTGTAAGA 276
      | | | | | | | | | |
Db     102  GAAACTATGTGAGA 89

```

RESULT 5

BE046932/c

LOCUS BE046932 484 bp mRNA linear EST 08-JUN-2000

DEFINITION hd92f06.x2 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2916995 3', mRNA sequence.

ACCESSION BE046932

VERSION BE046932.1 GI:8363985

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 484)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov

Seq primer: -40UP from Gibco

High quality sequence stop: 476.

FEATURES Location/Qualifiers

source 1..484

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2916995"

/tissue_type="pooled germ cell tumors"

/lab_host="DH10B"

/clone_lib="NCI_CGAP_GC6"

/note="Vector: pT7T3D-PacI; Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_GC4 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1257096-1258631, 1469064-1470983, and

1475592-1476743). Subtraction by Bento Soares and M.
Fatima Bonaldo. "

ORIGIN

Query Match 11.4%; Score 104.8; DB 2; Length 484;
Best Local Similarity 68.5%; Pred. No. 6.6e-18;
Matches 174; Conservative 0; Mismatches 77; Indels 3; Gaps 2;

```

Qy      24  TGTCATGGGGGTGCATGAGCAGCCCAGTGGAGAGGTGCACTTGGTGAGAAACCGATGCCT  83
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     338  TTTCATGAGGATACTCAAGCATTCTCTATGGAGAGATCCACATGGTGAGAAACTGAAGCCT  279

Qy      84  -CTGCCAACCACCTGCACTAACCTGCTGGGTCTGAGACTGAGCCACTTTGGAAGCTGATC  142
      | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     278  CCTACCAAGAGCCAGCACCAACTTGCCAGCTATGTGAATGAGCCATCTTAGAAGTGGGTT  219

Qy     143  TTGGAGCACCAAGTCAAGCCCTTAGCTGGCTGCAGCCACAGCCAACAACAAGACTGCAACC  202
      | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     218  CTCTAGCCCTAGTCAGGCCTTCATATGACTGCAGCCAGGGCTGATATTTTGACTACAACC  159

Qy     203  TCCTGGGGGATCCTGAGCCAGAATCCCCTGGCTAAATTGCTCCTTGATTCTTAACCCACA  262
      | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     158  TCATGAGAGA--CTGAGCCACAACAACCTAGCTAAGAAGCTCCTGAATTCCTACCAACA  101

Qy     263  GAAATTGTGTAAGA 276
      | | | | | | | | | |
Db     100  GAAACTATGTGAGA 87

```

RESULT 6

DB100179

LOCUS DB100179 500 bp mRNA linear EST 18-JAN-2008

DEFINITION DB100179 TESTI4 Homo sapiens cDNA clone TESTI4052031 5', mRNA
sequence.

ACCESSION DB100179

VERSION DB100179.1 GI:83431172

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 500)

AUTHORS Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.

TITLE Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes

JOURNAL Genome Res. 16 (1), 55-65 (2006)

PUBMED 16344560

COMMENT Contact: Takao Isogai
 FLJ Project (HRI Team)
 Helix Research Institute
 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
 Email: flj-cdna@nifty.com
 NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.

FEATURES Location/Qualifiers

source 1. .500
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="TESTI4052031"
 /tissue_type="testis"
 /clone_lib="TESTI4"
 /note="Vector: pME18SFL3"

ORIGIN

Query Match 11.3%; Score 104.4; DB 10; Length 500;
 Best Local Similarity 67.8%; Pred. No. 8.6e-18;
 Matches 175; Conservative 0; Mismatches 81; Indels 2; Gaps 2;

Qy	19	TGCTGTGTCATGGGGGTGCATGAGCAGCCCAGTGGAGAGGTGCACTTGGTGAGAAACCGA	78
Db	1	TACCATGTTGTGAGGCTGCTCAAGTAGCCTTGTGGAGAGGTCCACCTGGAGAGAAAACAA	60
Qy	79	TGCCTCTGCCAACCACTGCACTAACCTGCTGGGTCTGAGACTGAGCCACTTTGGAAGCT	138
Db	61	GGCCTTTGCC-ACAGCCGGCACCAACTTGCCAACCATCTGAGTCAGCCACCTTGGAAGTG	119
Qy	139	GATCTTGGAGCACCACTCAAGCCCTTAGCTGGCTGCAGCCACAGCCAACAACAAGACTGC	198
Db	120	GATCCTCCAGCCTCATTCCAGCCTTCAGATGACCACTGTCCCAGCTCATATCTAGACTGC	179
Qy	199	AACCTCCTGGGGGATCCTGAGCCAGAATCCCCTGGCTAAATTGCTCCTTGATTCTTAACC	258
Db	180	AACCTCATGAGAGACCTGGAGCCAGAA-CACCCAGCTAAGCTGTTCTGAATTCCCGACT	238
Qy	259	CACAGAAATTGTGTAAGA	276

Db 239 CACAGAGACTATGAAAGA 256

RESULT 7

CC492937/c

LOCUS CC492937 757 bp DNA linear GSS 17-JUN-2003

DEFINITION CH240_327A8.T7 CHORI-240 Bos taurus genomic clone CH240_327A8,
genomic survey sequence.

ACCESSION CC492937

VERSION CC492937.1 GI:31803851

KEYWORDS GSS.

SOURCE Bos taurus (cattle)

ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 757)

AUTHORS Holt,R., Stott,J., Yang,G., Barber,S., Smailus,D., Prabhu,A.-L.,
Tsai,M., Cloutier,A., Lee,D., Girn,N., Olson,T., Mayo,M.,
Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,
Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S.,
Schein,J., Marra,M., de Jong,P., McWilliam,S., Barris,W.,
Dalrymple,B.P. and Tellam,R.

TITLE Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398

JOURNAL Unpublished (2003)

COMMENT Other_GSSs: CH240_327A8.TARBAC13P2

Contact: Rob Holt

Sequencing

The British Columbia Cancer Agency Genome Science Centre

600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6

Tel: 604-877-6085

Fax: 604-877-6276

Email: rholt@bcgsc.ca

Clones are derived from the bovine BAC library CHORI-240

(<http://www.chori.org/bacpac/bovine240.htm>). For BAC library

availability, please contact Pieter de Jong (pdejong@mail.cho.org).

Clones may be purchased from BACPAC Resources

(http://www.chori.org/bacpac/ordering_information.htm). This work

was undertaken as part of the International Bovine BAC Mapping

Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the

British Columbia Genome Sciences Centre, Canada.

Plate: 327 row: A column: 8

Seq primer: T7

Class: BAC ends.

FEATURES Location/Qualifiers

source 1. .757

/organism="Bos taurus"

/mol_type="genomic DNA"

/strain="breed: Hereford"

Query Match 11.2%; Score 103.2; DB 20; Length 757;
Best Local Similarity 64.1%; Pred. No. 2.1e-17;
Matches 173; Conservative 0; Mismatches 93; Indels 4; Gaps 1;

```

RESULT 8
DB073121
LOCUS          DB073121          526 bp      mRNA      linear      EST 18-JAN-2008
DEFINITION     DB073121 TESTI4 Homo sapiens cDNA clone TESTI4016394 5', mRNA
                sequence.
ACCESSION      DB073121
VERSION        DB073121.1  GI:83564345
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
  ORGANISM     Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
                Catarrhini; Hominidae; Homo.
REFERENCE      1  (bases 1 to 526)

```

AUTHORS Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T., Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H., Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N., Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M., Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A., Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.

TITLE Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes

JOURNAL Genome Res. 16 (1), 55-65 (2006)

PUBMED 16344560

COMMENT Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.

FEATURES Location/Qualifiers

source 1. .526
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TESTI4016394"
/tissue_type="testis"
/clone_lib="TESTI4"
/note="Vector: pME18SFL3"

ORIGIN

Query Match 11.0%; Score 101.4; DB 10; Length 526;
Best Local Similarity 65.7%; Pred. No. 6.2e-17;
Matches 178; Conservative 0; Mismatches 91; Indels 2; Gaps 2;

Qy 7 GAGGGGAATGGCTGCTGTGTCATGGGGGTGCATGAGCAGCCCAGTGGAGAGGTGCACTTG 66
| ||| | | ||| ||| | || | ||| ||| ||| | || |

Db 10 GGGGGAAGTTGCTATGTTGTGAGGGCACTCATGCAGCCCATCAATGGTGAGATTCATGTG 69

Qy 67 GTGAGAAACCGATGCCT-CTGCCAACCACTGCACTAACCTGCTGGGTCTGAGACTGAGC 125
|| ||| || |||| ||||| ||| ||||| ||| | | || |||||

Db 70 CCAAGGAAGTGAAGCCTCCTGCCAATAACCAGCACTAACTTGCCAGACATATGAGTGAGC 129

Qy 126 CACTTTGGAAGCTGATCTTGGAGCACCAAGTCAAGCCCTTAGCTGGCTGCAGCCACAGCCA 185
|| | |||| ||| | ||| ||||| ||| | || ||| |||

Db 130 TACCTCAGAAGCAAATCCTCCAGCCCAGTCAAGCCTTCAGATGATTACAGCCCTGGCTA 189

```
Qy      186 ACAACAAGACTGCAACCTCCTGGGGGATCCTGAGCCAGAATCCCCTGGCTAAATTGCTCC 245
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      190 ATATCTTGACTG-AACTTCATGTGAGACCTTGAGCCAGAATCATCTAGCTAAGGCATTCC 248

Qy      246 TTGATTCTTAACCCACAGAAATTGTGTGAAGA 276
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      249 TGGATTTCCTGCCCCACAGCAATTATGTGAGA 279
```

RESULT 9

EI735387

LOCUS EI735387 669 bp DNA linear GSS 28-OCT-2007

DEFINITION CHORI51421012TR BAC library from the primary breast tumor sample
s104 Homo sapiens genomic clone CHORI514_21_012, genomic survey
sequence.

ACCESSION EI735387

VERSION EI735387.1 GI:158756676

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 669)

AUTHORS Raphael,B.J., Volik,S.V., Yu,P., Wu,C., Huang,G.Q., Waldman,F.,
Costello,J., Pienta,K., Mills,G., Bajsarowicz,K., Kobayashi,Y.,
Sridharan,S., Paris,P., Tao,Q.Z., Gray,J.W., Cheng,J.F., de
Jong,P., Nefedov,M., Aerni,S., Brown,R.P., Bashir,A.,
Padilla-Nash,H.M. and Collins,C.C.TITLE A sequence-based survey of the complex structural organization of
tumor genomes

JOURNAL Unpublished (2007)

COMMENT Contact: Volik SV
Colin Collins' lab
UCSF Department of Urology
UCSF Box 0808, San Francisco, CA 94143-0808, USA
Tel: 415 502 7067
Fax: 415 476 8218
Email: svolik@cc.ucsf.edu
Seq primer: KBR/TJ 5'CTGGCCGTCGACATTTAGG-3'.

FEATURES Location/Qualifiers

source 1. .669

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/clone="CHORI514_21_012"

/tissue_type="Primary tumor"

/clone_lib="BAC library from the primary breast tumor
sample s104"

/note="Organ: Breast; Vector: pBACGK1.1; Site_1: HindIII;
This library was constructed from s104 primary breast
tumor sample by Dr. M. Nefedov (Dr. Peter de Jong's
laboratory). The patient did not receive chemo- or
radiotherapy and did not have recurrence for 10 years
after surgery."

ORIGIN

Query Match 11.0%; Score 101.2; DB 22; Length 669;
Best Local Similarity 62.8%; Pred. No. 7.6e-17;
Matches 157; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

```

Qy      18 CTGCTGTGTCATGGGGGTGCATGAGCAGCCCAGTGGAGAGGTGCACTTGGTGAGAAACCG 77
      ||| | ||| | | | | | | | | | | | | | | | | | | |
Db     130 CTGCCATGTTCGTGAGAACACACAAGCAGCCCCAAGGAGGGGTCCGTGTGGTAAGGAAGT 189

Qy      78 ATGCCTCTGCCAACCACCTGCACTAACCTGCTGGGTCTGAGACTGAGCCACTTTGGAAGC 137
      | || ||| | | | | | | | | | | | | | | | |
Db     190 AGACCCCTGCTGGCAGCCAACACCACTTTGCCAACATTGTGAGTGACCCATCTTGGACAT 249

Qy     138 TGATCTTGAGCACCAGTCAAGCCCTTAGCTGGCTGCAGCCACAGCCAACAACAAGACTG 197
      | | | | | | | | | | | | | | | | | | | |
Db     250 AGGTTCTGACATCCTGTGAAGCCTTCAGCTGACTGCAGCCCTGTCTGACATCTTGACTG 309

Qy     198 CAACCTCCTGGGGGATCCTGAGCCAGAATCCCCTGGCTAAATTGCTCCTTGATTCTTAAC 257
      ||| ||| | | | | | | | | | | | | | | |
Db     310 CAACCTCATGAGAGATCCCAAACAAGAACTACCCAGCTAAGCCATTCCCTGATTCTGAT 369

Qy     258 CCACAGAAAT 267
      ||| ||| |||
Db     370 CCACAGAAAT 379

```

RESULT 10

BQ021722/c

LOCUS BQ021722 702 bp mRNA linear EST 27-MAR-2002
DEFINITION UI-H-DH1-axi-n-10-0-UI.s1 NCI_CGAP_DH1 Homo sapiens cDNA clone
IMAGE:5829321 3', mRNA sequence.
ACCESSION BQ021722
VERSION BQ021722.1 GI:19757000
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 702)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

JOURNAL

COMMENT

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. Jose Mercuende

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

The following repetitive elements were found in this cDNA sequence: 16-517, >MLT1F#LTR/MaLR (matched compliment)

Seq primer: M13 FORWARD

POLYA=Yes.

FEATURES

source

Location/Qualifiers

1. .702

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5829321"

/tissue_type="Metastatic Chondrosarcoma"

/dev_stage="Adult"

/lab_host="DH10B (Life Technologies)"

/clone_lib="NCI_CGAP_DH1"

/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; NCI_CGAP_DH1 is a normalized cDNA library containing the following tissue(s): VS-8 Cell line from Metastatic Chondrosarcoma in Lung. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AGATCATTGC.

TAG_TISSUE=lung

TAG_LIB=UI-H-DH1

TAG_SEQ=AGATCATTGC"

ORIGIN

Query Match

Best Local Similarity

Matches

11.0%;

Score 101.2;

DB 4;

Length 702;

62.8%;

Pred. No. 7.7e-17;

157;

Conservative

0;

Mismatches

93;

Indels

0;

Gaps

0;

Qy

18

CTGCTGTGTCATGGGGGTGCATGAGCAGCCCAGTGGAGAGGTGCACTTGGTGAGAAACCG

77

|||| |||| || | || ||||| ||| ||| ||| ||| |

Db 341 CTGCCATGTCGTGAGAACACACAAGCAGCCCCAAGGAGGGGTCCGTGTGGTAAGGAACTG 282

Qy 78 ATGCCTCTGCCAACCACCTGCACTAACCTGCTGGGTCTGAGACTGAGCCACTTTGGAAGC 137
| || |||| | || ||| | ||| || || ||| ||| ||||

Db 281 AGACCCCTGCTGGCAGCCAACACCACTTTGCCAACATTGTGAGTGACCCATCTTGGACAT 222

Qy 138 TGATCTTGGAGCACCAGTCAAGCCCTTAGCTGGCTGCAGCCACAGCCAACAACAAGACTG 197
| | || || || |||| | |||| ||||| | ||| | ||||

Db 221 AGGTTCCTGACATCCTGTGAAGCCTTCAGCTGACTGCAGCCCTGTCTGACATCTTGACTG 162

Qy 198 CAACCTCCTGGGGGATCCTGAGCCAGAATCCCCTGGCTAAATTGCTCCTTGATTCTTAAC 257
|||||| || | |||| | | |||| || |||| ||| ||||| | |

Db 161 CAACCTCATGAGAGATCCCAAACAAGAACTACCCAGCTAAGCCATTCCCTGATTCTTGAT 102

Qy 258 CCACAGAAAT 267
|||||||

Db 101 CCACAGAAAT 92

RESULT 11

DB074846

LOCUS DB074846 549 bp mRNA linear EST 18-JAN-2008

DEFINITION DB074846 TESTI4 Homo sapiens cDNA clone TESTI4018555 5', mRNA
sequence.

ACCESSION DB074846

VERSION DB074846.1 GI:83099624

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 549)

AUTHORS Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.TITLE Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes

JOURNAL Genome Res. 16 (1), 55-65 (2006)

PUBMED 16344560

COMMENT Contact: Takao Isogai

FLJ Project (HRI Team)

Helix Research Institute

2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan

Email: flj-cdna@nifty.com

NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.

```
FEATURES                      Location/Qualifiers
    source                      1. .549
                                /organism="Homo sapiens"
                                /mol_type="mRNA"
                                /db_xref="taxon:9606"
                                /clone="TESTI4018555"
                                /tissue_type="testis"
                                /clone_lib="TESTI4"
                                /note="Vector: pME18SFL3"
```

ORIGIN

```
Query Match          10.4%;   Score 95.4;   DB 10;   Length 549;
Best Local Similarity 63.1%;   Pred. No. 3.2e-15;
Matches 171;   Conservative 0;   Mismatches 86;   Indels 14;   Gaps 1;
```

```
Qy      17 GCTGCTGTGTCATGGGGGTGCATGAGCAGCCCAGTGGAGAGGTGCACTTGGTGAGAAACC 76
        |||||  |||||  |||  |||||  |||||  |||  |||||  |||
Db      7  GCTGCCCTGTCATGACGGTATGCATGCAGCCCCATGGAGAGGCCACATCATGAGGAACT 66

Qy      77 GATGCCTCTGCCAACCACCTGCACTAACCTGCTGGGTCTGAGACTGAGCCACTTTGGAAG 136
        |  |||||  |||  |||  |  |  |||||  |||
Db      67 AA-----CCAAGCACTAACTTGCCTGGTGTTCATAGGCCACTTTGGAGG 112

Qy     137 CTGATCTTGGAGCACCAGTCAAGCCCTTAGCTGGCTGCAGCCACAGCCAACAACAAGACT 196
        | |||  |  |||  ||||  |||  | ||  || |||||  ||  |  ||  ||||
Db     113 CAGATTCTCCAGCCTTGGTCAGGCCTTCAGATGACTGCAGCCCTGGCTGATGACTTGACT 172

Qy     197 GCAACCTCCTGGGGGATCCTGAGCCAGAATCCCCTGGCTAAATTGCTCCTTGATTCTTAA 256
        ||  |||||  ||  |  ||  |  |||||  |||  ||  ||  |||||  |||||  |
Db     173 GCCACCTCATGAGAGACCTTGAGCCAGAACCACCCAGCCAAGCCACTCCTGGATTCTGA 232

Qy     257 CCCACAGAAATTGTGTAAGACCTCCATCAGG 287
        |||||  ||  |  ||  ||  ||  |  |
Db     233 TCCACAGAACTAGGTGAAACGTCAAGAATG 263
```

RESULT 12

AW297623/c

```
LOCUS      AW297623                      446 bp      mRNA      linear      EST 16-JAN-2000
DEFINITION UI-H-BW0-ajg-f-08-0-UI.s1 NCI_CGAP_Sub6 Homo sapiens cDNA clone
            IMAGE:2731910 3', mRNA sequence.
ACCESSION  AW297623
VERSION    AW297623.1  GI:6704259
```

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 446)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html The following repetitive elements were found in this cDNA sequence: 49-424, >MLT1F#LTR/MaLR
Seq primer: M13 Forward
POLYA=Yes.

FEATURES Location/Qualifiers

source 1. .446
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2731910"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_Sub6"
/note="Vector: pT7T3D-PacI; Site_1: Not I; Site_2: Eco RI; NCI_CGAP_Sub6 is a subtracted library derived from BW, which consists of a mixture of four normalized libraries: NCI_CGAP_Brn50, NCI_CGAP_Lu13, NCI_CGAP_Ov18, GBC1 . The NCI_CGAP_Sub6 library had 7 million recombinants. A single-stranded DNA preparation of BW was used as a tracer in a subtractive hybridization with a driver comprising: the IMAGE pool (NCI_CGAP_Kid3 pool 1 LLAM 3334-3337, 3682-3683, 3798-3803 (IMAGE CloneIDs 1322376-1323911, 1456008-1456775, 1500552-1502855); NCI_CGAP_Kid5 pool 1 LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE CloneIDs 1323912-1325831, 1471368-1472903, 1492104-1493255); NCI_CGAP_Lu5 pool 1 LLAM 3575-3582, 3851-3854 (IMAGE CloneIDs 1414920-1417991, 1520904-1522439); NCI_CGAP_GC4 pool 1 LLAM 3164-3167, 3716-3720, 3733-3735 (IMAGE CloneIDs 1257096-1258631, 1469064-1470983, 1475592-1476743); NCI_CGAP_Pr22 pool 1 LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE CloneIDs 985608-986759,

1101192-1101959, 1217928-1220615); NCI_CGAP_Co10 pool 1
 LLAM 2644-2653, 2871-2872 (IMAGE CloneIDs
 1057416-1061255,1144584-1145351). (50% of the driver
 population), plus a pool of 3,840 arrayed clones from
 NCI_CGAP_Sub1 (IMAGE CloneIDs 2708616-2710535) and
 NCI_CGAP_Sub2 (IMAGE CloneIDs 2710536-2712455) (20% of
 the driver population), plus a pool of 11,136 clones from
 NCI_CGAP_Sub3 (IMAGE CloneIDs 2712456-2723591) (30% of
 the driver population). Subtraction was performed as
 previously described [Bonaldo, Lennon & Soares (1996):
 Normalization and Subtraction: Two Approaches To
 Facilitate Gene Discovery. Genome Research 6, 791-806.
 TAG_TISSUE=B cells germinal
 TAG_LIB=Gbc1
 TAG_SEQ=TCA"

ORIGIN

Query Match 10.3%; Score 94.8; DB 2; Length 446;
 Best Local Similarity 67.1%; Pred. No. 4.5e-15;
 Matches 149; Conservative 0; Mismatches 72; Indels 1; Gaps 1;

```

Qy      47 CCAGTGGAGAGGTGCACTTGGTGAGAAACCGATG-CCTCTGCCAACCACCTGCACTAACC 105
      || ||||| | || ||| || || || || ||||| | | || | ||
Db     328 CCTATGGAGATGCCCATGTGGGGAAGAATGGAGGTCCCCTGCCAACAATCAGCTTTGACA 269

Qy     106 TGCTGGGTCTGAGACTGAGCCACTTTGGAAGCTGATCTTGGAGCACCAGTCAAGCCCTTA 165
      ||| | ||| || ||||| | | ||| ||| ||| ||| ||| | |
Db     268 TGCCAAGCCTGTGAGGGAGCCATCTGGAAGGGAGTCTTCCAGCCCCAATCAGGCCTTCA 209

Qy     166 GCTGGCTGCAGCCACAGCCAACAACAAGACTGCAACCTCCTGGGGGATCCTGAGCCAGAA 225
      | || ||||| || ||| | | ||||| || | || | ||| |||
Db     208 GATGACTGCAGCCCTGGCCAGCTCCTTGACTGCAACCTCATGACAGACCTTGAGCAAGAA 149

Qy     226 TCCCCTGGCTAAATTGCTCCTTGATTCTTAACCCACAGAAAT 267
      | |||| ||| || ||||| |||||
Db     148 CTGCTCAGCTAAGCTGCTTCTGAATTCTTAACCTACAGAAAT 107

```

RESULT 13

AA535137/c

LOCUS AA535137 533 bp mRNA linear EST 21-AUG-1997
 DEFINITION nf87h12.s1 NCI_CGAP_Co3 Homo sapiens cDNA clone IMAGE:926951 3',
 mRNA sequence.
 ACCESSION AA535137
 VERSION AA535137.1 GI:2279390
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 533)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck,
M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 826 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 440.

FEATURES Location/Qualifiers
source 1. .533
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:926951"
/sex="pooled"
/tissue_type="colon"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Co3"
/note="Vector: pT7T3D-PacI; Site_1: Not I; Site_2: Eco RI;
1st strand cDNA was prepared from 12 pooled bulk tumor
samples and primed with a Not I - oligo(dT) primer.
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. "

ORIGIN

Query Match 10.3%; Score 94.6; DB 1; Length 533;
Best Local Similarity 62.3%; Pred. No. 5.3e-15;
Matches 175; Conservative 0; Mismatches 89; Indels 17; Gaps 1;

Qy 8 AGGGGAATGGCTGCTGTGTCATGGGGGTGCATGAGCAGCCCAGTGGAGAGGTGCACTTGG 67
|||| | | | |||| | || | ||||| |||| | | || ||
Db 478 AGGGAGAAGTCAGCTGTTCATGCGAGGGCACTCAAGCAGCCCTGTGGAAAGCTCTACGTGG 419
Qy 68 TGAGAAACCGATGCCTCTGCCAACCACCTGCACTAACCT-----GCTG 110
||||||| || |||| | | |||| || | || ||

Db	418	TGAGAAACTGAAGCCTCCTGCCAACACCAACAAGGAACTGAGACCTACTTCCAACAGCCA	359
Qy	111	GGTCTGAGACTGAGCCACTTTGGAAGCTGATCTTGGAGCACCAGTCAAGCCCTTAGCTGG	170
Db	358	TCTGAGTGATCGATCCATCCTGGAATCAGATCCTCCAGCCCCAGTCAAGTCTTCAGATGA	299
Qy	171	CTGCAGCCACAGCCAACAACAAGACTGCAACCTCCTGGGGGATCCTGAGCCAGAATCCCC	230
Db	298	CTGCAGCCCTGACCAGCATCTCAACTGCAACCTCGTGAATGACCCTGAGCCAGAACCATG	239
Qy	231	TGGCTAAATTGCTCCTTGATTCTTAACCCACAGAAATTGTG	271
Db	238	CAACTAAACTCTGCGTGGATTGATGACCCACAGAAACTGTG	198

RESULT	14
AW970418	
LOCUS	AW970418 637 bp mRNA linear EST 01-JUN-2000
DEFINITION	EST382499 MAGE resequences, MAGK Homo sapiens cDNA, mRNA sequence.
ACCESSION	AW970418
VERSION	AW970418.1 GI:8160263
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 637)
AUTHORS	Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and Quackenbush,J.
TITLE	Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray
JOURNAL	Unpublished (2000)
COMMENT	Contact: John Quackenbush The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 3528 Fax: 301 838 0208 Email: johnq@tigr.org Plate: 277 Seq primer: Forward.
FEATURES	Location/Qualifiers
source	1..637 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone_lib="MAGE resequences, MAGK" /note="Vector: pBluescriptSKm"

ORIGIN

Query Match 10.3%; Score 94.6; DB 2; Length 637;
 Best Local Similarity 62.3%; Pred. No. 5.6e-15;
 Matches 175; Conservative 0; Mismatches 89; Indels 17; Gaps 1;

```

Qy      8 AGGGGAATGGCTGCTGTGTCATGGGGGTGCATGAGCAGCCCAGTGGAGAGGTGCACTTGG 67
      |||| | | | |||| | || | ||||| |||| | | | || ||
Db     274 AGGGAGAAGTCAGCTGTCATGCGAGGGCACTCAAGCAGCCCTGTGGAAAGCTCTACGTGG 333

Qy     68 TGAGAAACCGATGCCTCTGCCAACCACCTGCACTAACCT-----GCTG 110
      ||||| || |||| | | |||| || | || |
Db    334 TGAGAAACTGAAGCCTCCTGCCAACACCAACAAGGAAGTGAAGACCTACTTCCAACAGCCA 393

Qy    111 GGTCTGAGACTGAGCCACTTTGGAAGCTGATCTTGGAGCACCAGTCAAGCCCTTAGCTGG 170
      | | || || || | |||| | |||| | || | ||||| || | | || |
Db    394 TCTGAGTGATCGATCCATCCTGGAATCAGATCCTCCAGCCCCAGTCAAGTCTTCAGATGA 453

Qy    171 CTGCAGCCACAGCCAACAACAAGACTGCAACCTCCTGGGGGATCCTGAGCCAGAATCCCC 230
      ||||| || || | ||||| || || | ||||| |||| |
Db    454 CTGCAGCCCTGACCAGCATCTCAACTGCAACCTCGTGAATGACCCTGAGCCAGAACCATG 513

Qy    231 TGGCTAAATTGCTCCTTGATTCTTAACCCACAGAAATTGTG 271
      |||| | | |||| | ||||| |||| |
Db    514 CAACTAAACTTTGCGTGGATTTCATGACCCACAGAAACTGTG 554
  
```

RESULT 15

CE539857/c

LOCUS CE539857 429 bp DNA linear GSS 28-SEP-2003
 DEFINITION tigr-gss-dog-17000366002738 Dog Library Canis lupus familiaris
 genomic, genomic survey sequence.
 ACCESSION CE539857
 VERSION CE539857.1 GI:36856638
 KEYWORDS GSS.
 SOURCE Canis lupus familiaris (dog)
 ORGANISM Canis lupus familiaris
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Carnivora; Caniformia; Canidae;
 Canis.
 REFERENCE 1 (bases 1 to 429)
 AUTHORS Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
 Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
 Venter,J.C.
 TITLE The dog genome: survey sequencing and comparative analysis
 JOURNAL Science 301 (5641), 1898-1903 (2003)
 PUBMED 14512627
 COMMENT Contact: Kirkness EF
 The Institute for Genomic Research

Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.

FEATURES Location/Qualifiers
 source 1. .429
 /organism="Canis lupus familiaris"
 /mol_type="genomic DNA"
 /strain="Standard Poodle"
 /sub_species="familiaris"
 /db_xref="taxon:9615"
 /clone_lib="Dog Library"
 /note="Site_1: BstXI; Libraries were prepared from
 peripheral blood"

ORIGIN

Query Match 10.2%; Score 94.2; DB 20; Length 429;
Best Local Similarity 63.4%; Pred. No. 6.5e-15;
Matches 144; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

```
Qy      17  GCTGCTGTGTCATGGGGGTGCATGAGCAGCCAGTGGAGAGGTGCACTTGGTGAGAAACC 76
          |||||  |||||  || | |  | | |  |||||  || || | |||  |||||
Db      312 GCTGCCATGTCAATAGGATCCTCAAAGACTCTTGTGGAGAAGTTCACATGGTAAGAAACT 253

Qy      77  GATGCCTCTGCCAACCACCTGCACTAACCTGCTGGGTCTGAGACTGAGCCACTTTGGAAG 136
          || || |  | | |||| |||  |  |||  ||| || |||||  |||||
Db      252 GAGGCTTAATGCTAACACCAGCATCAGTTTGCCAACACTGGGAGTGAGCTCTCTTGGAAG 193

Qy     137  CTGATCTTGGAGCACCAGTCAAGCCCTTAGCTGGCTGCAGCCACAGCCAACAACAAGACT 196
          ||| |  || | ||||| || | |  |  ||||  ||| ||| |  ||||
Db     192  TGAATCCTTTAGGCCAAGTCAAGCCTTTGGATAACCATAGCCCTGGCCCACATCTTGACT 133

Qy     197  GCAACCTCCTGGGGGATCCTGAGCCAGAATCCCCTGGCTAAATTGCT 243
          |||||  || || |  || ||||| ||||| | ||  || ||  ||||
Db     132  GCAACTTCATGAGAGACCCTGAGCCAGAACCACCCAGCAAAGCTGCT 86
```

Search completed: May 31, 2009, 22:23:27
Job time : 1892 secs

SCORE 3.0